

I enclose a file containing the nucleotide sequence of the PRE7(IAP) as is the provisional name we have selected for this element, which is involved in posttranscriptional regulation.

We have identified elements that have homology to PRE7(IAP) from the databases. The alignment of these elements is provided. None of these elements has been characterized functionally. Our work is the first that identified these sequences as potential posttranscriptional control elements. At present, PRE7(IAP) is the only one for which we have functional data.

In the aligned sequences, our functional PRE7(IAP) is called FNC3B. Only the "core" 231 nt element containing the functional element is shown.

FNC3B

GTGGGGTGCAGGCTAAGCACTGCACAGAGGATAGCTTGCTGTTGGCATCCTGTGGAAGG
CACGTCTGATTGCATGAAGGTTCAAGTGTCTAGTTCCCTTCCCCCAGGAAAAACGACACG
GGAGCTGGCCAAGACCTCTCTGGGTGATGAGCCTAAGGGATGGTTTTGTGTAGGGCCCCCT
ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT

PRE7 (IAP)

CTTTCGCCATGGTAGCATAGGCTTTTGCTGCAGTGGAGGCGGGACAATCTCCTCAGATTC
GGTTTGCCGCTCTAAAAGAAATTATGCTGCGTTATGCCGTGGGGTGCAGAGGCTAAGCACT
GCACAGAGGATAGCTTGCTGTTGGCATCCTGTGGAAGGCACGTCTGATTGCATGAAGGTT
CAGTGTCTAGTTCCCTTCCCCCAGGAAAAACGACACGGGAGCTGGCCAAGACCTCTCTG
GGTGATGAGCCTAAGGGATGGTTTTGTGTAGGGCCCCCTATGCTTGCACACTGGGGATCAG
ACCTCTACCTTCACCCATGAGGCTTGCTTGCAGCAATTAAGATCTGGCCATAGGTTAATT
AACATCCTGGCCTTTTGATGCACCTGCCACAAG

the underlined sequences correspond to the aligned fragment
FNC3B

The alignment of the sequences found in the database
follows:

FNC3B.seq Begin:99 End:329
GB:X57268 Begin:1599 End:1841
GB:M10134 Begin:7308 End:7538
GB:X01172 Begin:2423 End:2668
GB:M12515 Begin:188 End:418
AE000664.seq Begin:118019 End:118264
GB:M18252 Begin:3418 End:3662
GB:M18251 Begin:1630 End:1874
GB:S74315 Begin:3824 End:4068
GB:M10062 Begin:2711 End:2955
GB:E00593 Begin:2711 End:2955
E01116.gb_pat Begin:2711 End:2955
GB:E00594 Begin:3537 End:3781
E01117.gb_pat Begin:4388 End:4632
GB:X54077 Begin:2200 End:2444
GB:X04120 Begin:4477 End:4721
GB:X97915 Begin:839 End:1083
GB:M17551 Begin:6474 End:6721
GB:U58494 Begin:6192 End:6439
U70139.gcg Begin:4321 End:4565

!!NA MULTIPLE ALIGNMENT 1.0
PileUp of: @CTE-setClg.list

Symbol comparison table: GenRunData:pileupdna.cmp CompCheck: 6876

GapWeight: 2
GapLengthWeight: 1

CTE-setClg.msf MSF: 251 Type: N 10:03 Check: 7329 ..

Name: FNC3B	Len: 251	Check: 6699	Weight: 1.00
Name: X57268	Len: 251	Check: 7888	Weight: 1.00
Name: M10134	Len: 251	Check: 8724	Weight: 1.00
Name: X01172	Len: 251	Check: 3612	Weight: 1.00
Name: M12515	Len: 251	Check: 112	Weight: 1.00
Name: AE000664	Len: 251	Check: 4675	Weight: 1.00
Name: M18252	Len: 251	Check: 1995	Weight: 1.00
Name: M18251	Len: 251	Check: 2423	Weight: 1.00
Name: S74315	Len: 251	Check: 1833	Weight: 1.00
Name: M10062	Len: 251	Check: 1751	Weight: 1.00
Name: E00593	Len: 251	Check: 1751	Weight: 1.00
Name: E01116	Len: 251	Check: 1751	Weight: 1.00
Name: E00594	Len: 251	Check: 1811	Weight: 1.00
Name: E01117	Len: 251	Check: 1811	Weight: 1.00
Name: X54077	Len: 251	Check: 2537	Weight: 1.00
Name: X04120	Len: 251	Check: 2537	Weight: 1.00
Name: X97915	Len: 251	Check: 3213	Weight: 1.00
Name: M17551	Len: 251	Check: 5139	Weight: 1.00
Name: U58494	Len: 251	Check: 5139	Weight: 1.00
Name: U70139	Len: 251	Check: 1928	Weight: 1.00

//

1

FNC3B	GTGGGGTGCG	AGGCTAAGC.	ACTGCACAGA	GGATAGCTT.	...GCTGT.T	GG.CATCCTG	T.GGAAGGCA	CGTC
X57268	GAGAGTTGTA	AGACTAAGT.	ACTGCACAGA	GATTAGTCTA	GAAGCTGT.T	GGACAGTCTC	T.GAGAGGCA	TGTC
M10134	AGGAGTTGCA	AGGCTAAGC.	ACTGCACAGG	AGAGG.TCTG	CGG..TATAA	CGACTTTCTC	CTGGGAGATA	AGTC
X01172	GAGAGTTGCA	CGGCTAAGC.	ACTGCAGTAG	AAGGGCTCTG	CGGCACATAT	GAGCCTATTC	TAGGGAGACA	TGTC
M12515	GAGAGTCG.A	AGGCTAAGCA	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAGACA	TGTC
AE000664	GAGAGTTGCA	CGGCTAAGC.	ACTGCAGTAG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAGACA	TGTC
M18252	GAGAGTTGCA	CGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAGACA	TGTC
M18251	GAGAGTTGCA	AGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAGACA	TGTC
S74315	GAGAGTTGCA	CGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAGACA	TGTC
M10062	GAGAGTTGCA	AGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAAACA	TGTC
E00593	GAGAGTTGCA	AGGCTAAGC.	ACTGCAATGG	AAAGGCTCTG	CGGCATATAT	GAGCCTATTC	TAGGGAAACA	TGTC

```
E01116 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAAACA TGTC
E00594 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAGACA TGTC
E01117 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAGACA TGTC
X54077 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAGACA TGTC
X04120 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAGACA TGTC
X97915 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GTGCCTATTC TAGGGGAGACA TGTC
M17551 GAGAGTTGCA CGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAGACA TGTC
U58494 GAGAGTTGCA CGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGAGACA TGTC
U70139 GAGAGTTGCA AGGCTAAGC. ACTGCAATGG AAAGGCTCTG CGGCATATAT GAGCCTATTC TAGGGNGACA TGTC

101
FNC3B AGTTCCCTTC .CCCCAGGAA AAACGACACG GGAGCTGGCC AAGACC.TCT CTGGGTGA.. ....TGAGC CT.A
X57268 AGGGACCTTT .CCCCAGAAA AAAGGGCACA GGAGCAGGTC AGGGTT.ACT CTGGGTAAAG ATCTGTGGGC CT.G
M10134 GATCTCCTT. .CCCCAGAAA AAAAGACATC GGA. CTGGTC AGGACTTCCT CTGGGGATAA G.....ACC CTGG
X01172 AGTGTCTTTC TCCCCAGGAA AAACGGCAGC GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTGAGC CT.A
M12515 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGA. CAGGTC AGGGTT.GCT CTGGGTAAAA .CCTGTAAGC CT..
AE000664 AGTGTCTTTC TCTCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTGAGC CT.A
M18252 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTGAGC CT.A
M18251 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GTT CTGGGTAAAA GCCTGTAAGC CT.A
S74315 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.A
M10062 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.A
E00593 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.A
E01116 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.A
E00594 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.A
E01117 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.A
X54077 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTGAGC CT.A
X04120 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTGAGC CT.A
X97915 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTGAGC CT.A
M17551 AGTTGCCCT. TCCCCAGGCA AAACGACAC. GGAGCAGGTC AGGGTTGGCT CTGGGTAAAA GCCTGTGAGC CTCF
U58494 AGTTGCCCT. TCCCCAGGCA AAACGACAC. GGAGCAGGTC AGGGTTGGCT CTGGGTAAAA GCCTGTGAGC CTCF
U70139 AGTTCCCTTC .CCCCAGGCA AAACGACACG GGAGCAGGTC AGGGTT.GCT CTGGGTAAAA GCCTGTAAGC CT.F

201
FNC3B TG.CTTGCAC ACTGGGGATC AGACCTCTAC CTTACCCCAT GAGGCTTGCT T
X57268 AC.ATGACAC ACTGGGGATC AGACCTCTAC CTCTACCCAC GGAGCTTGCT T
M10134 T...TTGCAC A.TGGGGATT TGACCTCTAT CTCCACTC.C AAAGTTGTGG G
X01172 T.ACCTGCAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
M12515 TTACCTACAC ACTGGG... TGACCTCTAT CT.CACTCTC ATCAATATGG T
AE000664 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAGTTGGG T
M18252 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
M18251 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
S74315 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCAGTCTC ATTAATATGG G
M10062 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
E00593 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
E01116 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
E00594 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
E01117 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
X54077 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
X04120 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
X97915 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
M17551 TGACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
U58494 TGACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G
U70139 T.ACCTACAC ACTGGGGATT TGACCTCTAT CTCCACTCTC ATTAATATGG G

251
```